



A Land and Ocean Microwave Cloud Classification Algorithm Derived from AMSU-A and -B, Trained Using MSG-SEVIRI Infrared and Visible Observations

Filipe Aires, Francis Marquisseau, Catherine Prigent, Genevieve Seze

► To cite this version:

Filipe Aires, Francis Marquisseau, Catherine Prigent, Genevieve Seze. A Land and Ocean Microwave Cloud Classification Algorithm Derived from AMSU-A and -B, Trained Using MSG-SEVIRI Infrared and Visible Observations. Monthly Weather Review, 2011, 139 (8), pp.2347-2366. 10.1175/mwr-d-10-05012.1 . hal-01117322

HAL Id: hal-01117322

<https://hal.science/hal-01117322>

Submitted on 16 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Land and Ocean Microwave Cloud Classification Algorithm Derived from AMSU-A and -B, Trained Using MSG-SEVIRI Infrared and Visible Observations

FILIPPE AIRES* AND FRANCIS MARQUISSEAU

Laboratoire de Météorologie Dynamique/IPSL/CNRS, Université de Paris VI, Jussieu, France

CATHERINE PRIGENT

Laboratoire de l'Etude du Rayonnement et de la Matière en Astrophysique, CNRS, Observatoire de Paris, Paris, France

GENEVIÈVE SÈZE

Laboratoire de Météorologie Dynamique/IPSL/CNRS, Université de Paris VI, Jussieu, France

(Manuscript received 12 October 2010, in final form 3 February 2011)

ABSTRACT

A statistical cloud classification and cloud mask algorithm is developed based on Advanced Microwave Sounding Unit (AMSU-A and -B) microwave (MW) observations. The visible and infrared data from the Meteosat Third Generation-Spinning Enhanced Visible and Infrared Imager (MSG-SEVIRI) are used to train the microwave classifier. The goal of the MW algorithms is not to fully reproduce this MSG-SEVIRI cloud classification, as the MW observations do not have enough information on clouds to reach this level of precision. The objective is instead to obtain a stand-alone MW cloud mask and classification algorithm that can be used efficiently in forthcoming retrieval schemes of surface or atmospheric parameters from microwave satellite observations. This is an important tool over both ocean and land since the assimilation of the MW observations in the operational centers is independent from the other satellite observations.

Clear sky and low, medium, and opaque-high clouds can be retrieved over ocean and land at a confidence level of more than 80%. An information content analysis shows that AMSU-B provides significant information over both land and ocean, especially for the classification of medium and high clouds, whereas AMSU-A is more efficient over ocean when discriminating clear situations and low clouds.

1. Introduction

Cloud detection and classification are generally based on satellite visible and infrared observations. First, these wavelength domains show a very high degree of sensitivity to the presence of clouds, with good spatial resolution. Second, these observations are available from

geostationary satellites, making it possible to track cloud developments with frequent measurements of the same locations. For instance, visible and thermal infrared observations from polar orbiters and geostationary meteorological satellites have been combined by the International Satellite Cloud Climatology Project (ISCCP) to provide global cloud information dataset since 1983, every 3 h, with an ~ 30 -km spatial resolution (Rossow and Schiffer 1999). To obtain global uniformity in the ISCCP climatology, which is built from observations from several instruments, the number of channels has been limited to one visible (VIS) and one infrared (IR) window. A serial threshold technique is used to detect clouds; furthermore, a radiative transfer code and ancillary data such as temperature and humidity profiles are used to retrieve the cloud-top pressure and cloud optical thickness (Rossow et al. 1985).

* Additional affiliation: Laboratoire de l'Etude du Rayonnement et de la Matière en Astrophysique, CNRS, Observatoire de Paris, Paris, France.

Corresponding author address: F. Aires, CNRS/IPSL/Laboratoire de Météorologie Dynamique, Université Pierre et Marie Curie, Case 99, 4, Place Jussieu, F-75252 Paris, CEDEX 05, France.
E-mail: filipe.aires@lmd.jussieu.fr

The information obtained from additional VIS and/or IR channels can improve the cloud detection, the separation of broken clouds from thin cirrus, and/or the differentiation of multilayered clouds from midlevel clouds. This has been pointed out by many authors (Saunders 1986; Desbois et al. 1982; Coakley 1983; Inoue 1985; Baum et al. 1997; Jin and Rossow 1997). Different approaches have used the information from these spectral bands, for example, physically based threshold techniques (Derrien and Gléau 2005, 2010), clustering techniques without a priori knowledge (Sèze and Pawlowska 2001; Ambroise et al. 2000), and neuronal and fuzzy logic approaches (Baum et al. 1997; Miller and Emery 1997). Recently, the multispectral capabilities of the new generation of VIS and IR imagers, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the *Terra* and *Aqua* platforms or the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the Meteosat geostationary satellite have been investigated. These new observations have shown large improvements in their cloud detection and the cloud property retrieval (Ackerman et al. 1998; Frey et al. 2008; Derrien and Gléau 2010). Furthermore, active measurements from the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument lidar on board the Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) platform of the A-Train constellation have been used to evaluate the VIS and IR passive measurements algorithms (Holz et al. 2008; Sèze et al. 2009).

Microwave observations are less sensitive to thin clouds than visible or infrared measurements. However, contrarily to visible and infrared observations, which only sense radiation scattered or emitted from the top of the clouds, microwave radiation can propagate through clouds. As a consequence, microwave observations have a better ability to sense the total cloud layer and have the potential to estimate cloud water and ice contents. At frequencies below ~ 80 GHz, the microwave signal is essentially dominated by emission and absorption by liquid clouds and rain and is little affected by the presence of ice. At higher frequencies, the scattering effect on frozen particles increases. Ice particles modify the upwelling radiation by scattering photons away from the satellite sensors, causing a brightness temperature depression. Over ocean, cloud liquid water paths are routinely estimated from the cloud emissions measured between 19 and 85 GHz by imagers such as the Special Sensor Microwave Imager (SSM/I) or the Advanced Microwave Scanning Radiometer (AMSR) (Alishouse et al. 1990; Greenwald et al. 1993; Ferraro et al. 1996; O'Dell et al. 2008). Over land, the problem is more complicated. The land surface emissivity is usually close to unity, making atmospheric features difficult to identify

against such a background because of the limited contrast. In addition, the land surface emissivity is variable in space and time and difficult to model. However, efforts have been made to estimate cloud liquid water over land, using a priori information on the surface properties (Aires et al. 2001). From observations above ~ 80 GHz, cloud ice information has been extracted, from both imagers such as SSM/I and water vapor sounders such as the Advanced Microwave Sounding Unit-B (AMSU-B) (Greenwald and Christopher 2002; Hong and Heygster 2005; Weng et al. 2003).

Passive microwave sounders such as AMSU-A and -B are operationally used to estimate temperature and water vapor atmospheric profiles (Radnoti et al. 2010). The quality of the retrieval partly depends on the knowledge of cloud presence and nature (Engelen and Stephens 1999). Cloud information could be extracted from visible and infrared observations. However, this implies that the observations are coincident in time and space, which is not often practical, especially in operational assimilation systems with sequential assimilation of the different instruments. The numerical weather prediction centers do not use IR–VIS information for the assimilation of MW observations.

The objective of this study is to develop a cloud classification directly from the microwave observations. The idea is to capitalize on the capacities of the new generation of geostationary instruments (e.g., MSG-SEVIRI) to precisely characterize the cloud presence and types. The Satellite Application Facility for Nowcasting (SAFNWC; information online at <http://nwcsaf.inm.es/>) cloud classification that is described in the next section will be used here (Derrien and Gléau 2005, 2010). This cloud classification has been extensively validated. Collocated MSG-SEVIRI and AMSU-A and -B observations will then help train a microwave-only cloud classification. This stand-alone MW cloud classification and cloud mask algorithm will be designed for use over both ocean and land. Snowy data are not considered in this study, the geostationary satellite coverage of the snow-covered regions is not ideal due to its limited coverage of the high latitudes. The MW classifier cannot obtain a similar degree of precision or level of detail as the SEVIRI algorithm; in particular, it would be difficult to treat multilevel clouds or partially cloud-covered scenes. However, the fact that some cloud signals cannot be measured by the MW observations means that they have a low impact on them, implying that they would have a limited impact on the forthcoming retrieval scheme. Furthermore, the MW classifier works at a coarser horizontal resolution than does the SEVIRI algorithm. This implies that some averaging appears in the cloud properties as seen by the MW instrument.

The satellite datasets are described in section 2. Three classification methods are tested in section 3. The classification results are presented in section 4, with special emphasis on the analysis of the microwave information content. Finally, section 5 concludes this work.

2. The datasets

a. The SAFNWC cloud classification from MSG-SEVIRI

The European Space Agency (ESA) and the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) have joined efforts to develop the Meteosat Second Generation (MSG) mission (Schmetz et al. 2002). This project capitalizes on the first generation of spacecraft in the Meteosat weather satellite series but technical innovations have boosted its performance. The first satellite of the series, *Meteosat-8*, was launched in 2003 and entered into service at EUMETSAT in early 2004. The second satellite was launched in 2005 and became *Meteosat-9* in July 2006.

The main instrument on MSG is SEVIRI (Govaerts et al. 2001). This optical imaging radiometer observes the earth's atmosphere and surface at 12 different wavelengths (compared to the 3 channels of the previous generation): 4 visible–near-infrared (NIR) channels (0.4–1.6 μm) and 8 IR channels (3.9–13.4 μm). A very important feature of SEVIRI is its continuous imaging of the earth, with a baseline repeat cycle of 15 min (compared to 30 min for previous instruments). The imaging sampling distance is 3 km \times 3 km at the subsatellite point (about 6 km \times 6 km at the border of the analysis domain) for standard channels, and down to 1 km for the high-resolution visible (HRV) channel.

The two first steps of the SAFNWC algorithm, cloud detection and classification, rely on multispectral threshold tests applied at the pixel scale to a set of spectral and textural features (Derrien and Gléau 2005). The left column of Table 1 provides the names of the SEVIRI cloud classes. An important feature of this algorithm is that the threshold values depend on the illumination and viewing geometry and the geographical location. They are computed using ancillary data fields that feed a radiative transfer model [i.e., the radiative transfer model for the Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS), RTTOV]. The ancillary data are composed of atlases (i.e., height map and land–sea mask), climatological maps of sea surface temperature and continental reflectance, and numerical weather predictions of surface temperature, integrated atmospheric precipitable water, and atmospheric temperature and humidity profiles.

After isolating clear-sky from cloud-contaminated pixels, cloudy pixels are classified into two sets: 1) fractional

TABLE 1. SEVIRI cloud classes, and corresponding MW four-nebulosity classes used in this study. Boldface font signify SEVIRI classes kept in the new MW cloud classes.

Original cloud class	SEVIRI class description	New MW cloud class
0	Not processed	
1	Cloud-free land	Clear
2	Cloud-free sea	Clear
3	Land contaminated by snow	
4	Sea contaminated by snow/ice	
5	Very low and cumuliform clouds	Low
6	Very low and stratiform clouds	
7	Low and cumuliform clouds	
8	Low and stratiform clouds	Low
9	Medium and cumuliform clouds	Medium
10	Medium and stratiform clouds	Medium
11	High opaque and cumuliform clouds	
12	High opaque and stratiform clouds	High
13	Very high opaque and cumuliform clouds	
14	Very high opaque and stratiform clouds	High
15	High semitransparent thin clouds	
16	High semitransparent meanly thick clouds	
17	High semitransparent thick clouds	
18	High semitransparent above low or medium clouds	
19	Fractional clouds (subpixel water clouds)	
20	Undefined	

cloud and high semitransparent cloud and 2) low, medium, and high thick clouds. A separation between fractional, high semitransparent cloud (single layer) and high semitransparent cloud in a multilayered system is performed. For thick clouds, ancillary temperature and humidity profiles help discriminate between low, middle, and high clouds. This cloud classification scheme aims at retrieving 11 cloud types: clear, very low, low, midlevel, high, very high, very thin cirrus, cirrus, thick cirrus, cirrus over another layer, and partly covered pixels.

It is very difficult to reproduce from microwave observations the level of detail in a cloud classification that is obtained from instruments such as SEVIRI. It was thus decided to 1) keep only the clear and the opaque monolayer cases, 2) eliminate the SEVIRI-derived ambiguous cloud type, and 3) group the remaining original SEVIRI classes into four broad classes: clear sky, low clouds, medium cloud, and high cloud opaque. The correspondence of the SEVIRI and MW cloud classes is provided in Table 1. Since the SEVIRI algorithm is based on an empirical approach using channel differences, the precise determination of the height of each of the cloud classes is not possible.

The goal of the MW cloud classification will be to reproduce these four cloud types. Figure 1 provides an example of such cloud classification for 1 July 2006.

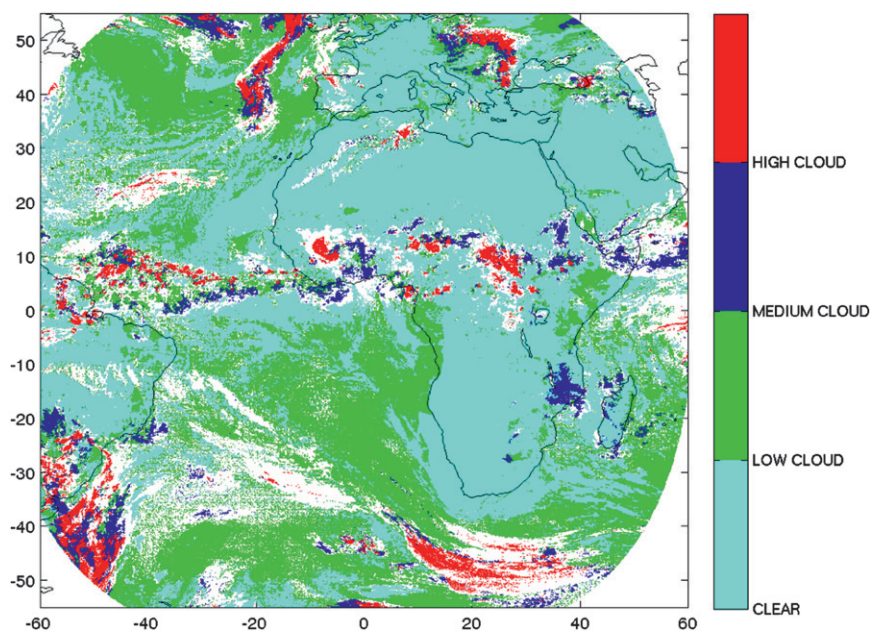


FIG. 1. SEVIRI cloud classification for 1 Jul 2006 in the clear sky and low-, medium-, and high-cloud classes.

The white areas are for the ambiguous cloud type situations.

At the time of this study, the available data (SEVIRI observations along with the derived cloud classification) covered June–October 2006. These five months will be used.

b. AMSU-A and -B on NOAA-16

Among other instruments, the *National Oceanic and Atmospheric Administration-16* (NOAA-16) satellite

includes the two MW instruments of interest in this study: AMSU-A and -B. The general characteristics of both instruments are provided in Table 2.

The AMSU-A measures the MW radiation for the retrieval of atmospheric temperature profiles, with 12 sounding channels between the 50- and 60-GHz O_2 bands, and three other channels at 23.8, 31.4, and 89 GHz. It is a cross-track scanning radiometer, with $\pm 48.3^\circ$ from nadir with a total of 30 earth fields of view of 3.3° per scan line, providing a nominal spatial resolution of 48 km at nadir.

TABLE 2. General instrument characteristics for AMSU-A and -B.

Instrument	AMSU-A		AMSU-B	
Scanning	Cross track		Cross track	
Spatial resolution	48 km (at nadir)		16 km (at nadir)	
Channel frequencies, in GHz (noise in K)	23.8 (0.30)		89.0	(0.37)
	31.4	(0.30)	157.0	(0.84)
	50.3	(0.40)	183.31 ± 1	(1.06)
	52.8	(0.25)	183.31 ± 3	(0.70)
	53.596 ± 0.115	(0.25)	183.31 ± 7	(0.60)
	54.4	(0.25)		
	54.94	(0.25)		
	55.5	(0.25)		
	$57.290\ 344 (=F_{LO})^*$	(0.25)		
	$F_{LO} \pm 0.217$	(0.40)		
	$F_{LO} \pm 0.3222 \pm 0.048$	(0.40)		
	$F_{LO} \pm 0.3222 \pm 0.022$	(0.60)		
	$F_{LO} \pm 0.3222 \pm 0.010$	(0.80)		
	$F_{LO} \pm 0.3222 \pm 0.0045$	(1.20)		
	89.0	(0.50)		

* Frequency of Local Oscillator (FLO).

The swath is approximately 2000 km and the instrument realizes one scan in 8 s.

The AMSU-B microwave radiometer is designed to measure the atmospheric water vapor profile, with three channels in the H_2O line at 183.31 GHz plus two window channels at 89 and 150 GHz that enable deeper penetration through the atmosphere down to the earth's surface. Each swath is made up of 90 contiguous individual pixels and scanned every 2.67 s.

AMSU-B pixels at nadir have a diameter of approximately 16 km and each cross-track pass has 90 of them. AMSU-B has the same swath width that AMSU-A does, but the measurements are done in one-third of the time, allowing for a good synchronization of both instruments. A 3×3 AMSU-B pixel matrix covers each of the AMSU-A pixels in order to facilitate the synergetic use.

c. Coupling of the AMSU observations with the SEVIRI-derived cloud classification

To use the datasets jointly, the first step consists of collocating them in space and time. The AMSU-B pixels are, by design, very easy to map onto AMSU-A pixels. The 3×3 AMSU-B pixels mentioned in section 2b are simply averaged and then paired to the associated AMSU-A pixel. A control step is run to suppress any missing scan lines. The SEVIRI observations with their centers inside an AMSU-A pixel are then associated with it. The spatial resolution of SEVIRI being much higher than that of AMSU, an ensemble of different SEVIRI cloud classes can be linked to a single AMSU pixel. How should one go about mapping these multiple SEVIRI classes onto each microwave pixel? A lot of work has been done in this regard. Mixed cloud categories or combined cloud types could have been considered but the dataset built here will train the MW cloud classifiers and there is not enough information in the MW observations to treat these complex cases. It is then recommended to limit, as much as possible, the number of ambiguous points (the heterogeneity of the SEVIRI cloud classification inside the bigger MW pixels will be considered in section 3c). A few rules are adopted to filter out these heterogeneous pixels:

- An AMSU pixel is labeled clear only if more than 95% of the associated SEVIRI classes are clear (different thresholds were tested and this number was found to be a good compromise).
- An AMSU pixel is labeled low (medium or high) cloud only if more than 80% of the associated SEVIRI classes are low (medium and high).
- All other AMSU pixels are excluded from the training dataset.

Table 1 provides a correspondence between the SEVIRI and our microwave cloud classifications.

The purpose of excluding some of the cloud situations in the training dataset is not to avoid processing these cases. The goal is to train the classification models with, as much as possible, "clean" situations so that the classifier can infer, from this training, unambiguous spectral signatures. After the training stage, the classifier can be used on ambiguous situations and it is expected to estimate the ambiguity level. Some of the clouds that are excluded in the training dataset such as multilevel clouds and AMSU pixels that are not homogeneous enough can impact, if they are thick enough, the MW measurements and therefore the retrievals. The classification algorithm is able to detect those clouds, although maybe not able to correctly classify them. The tests outlined in sections 4c and 4d are performed on nonfiltered scenes.

This collocation process allows for the construction of a dataset of AMSU-A and -B observations classified a priori, using the SEVIRI-derived cloud classification over ocean and various types of continental surfaces. The SEVIRI-derived cloud classification is used as the reference for the microwave cloud mask and classification. The microwave cloud classification will be "trained" on the SEVIRI classification.

d. Microwave surface emissivity information

As discussed in the introduction, microwave observations from AMSU-A and -B over land can include a significant contribution from the surface that will contaminate the cloud analysis. To facilitate the cloud classification, different surface types are considered. Instead of using external vegetation-soil information, a surface classification based on microwave emissivities has been adopted (Prigent et al. 2008). It is derived from a monthly mean emissivity climatology calculated from a decadal (1992–2001) SSM/I emissivity database (Prigent et al. 2006), a snow flag from the National Snow and Ice Data Center (NSDC), and a river, lake, and wetland (Prigent et al. 2007) dataset. An unsupervised clustering technique is applied to this emissivity climatology for the seven SSM/I channels. A Kohonen topological map (Prigent et al. 2001) is used to classify all the surface pixels. The clustering method is applied twice as follows: once for the snow-free pixels, then for the snow-covered pixels (the snow and ice information is extracted from the National Snow and Ice Data Center; ice pixels are not considered).

This emissivity-derived surface information ensures that each class represents a different pattern of behavior in terms of microwave emissivities and that the set of emissivity classes describes their full variability. The cloud classification results will be analyzed for each surface type. Table 3 lists the 10 surface types. Five classes are isolated for the snow-free regions, corresponding to

TABLE 3. Surface-type classes from 1) the NSIDC flag; (2) a lake, river, and wetland dataset, and (3) SSM/I microwave emissivities.

Class No.	Description
1	Arid
2	Low vegetation
3	Medium vegetation
4	Medium/high vegetation
5	High vegetation
6	Snow/ice type 1
7	Snow/ice type 2
8	Snow/ice type 3
9	Snow/ice type 4
10	Including standing water

vegetation densities, from dense forest (class 1) to desert surfaces (class 5). Four snow classes are also determined (classes 6–9). Pixels with more than 10% standing water are not considered in the clustering scheme and are grouped into class 10. This class includes areas of rivers or lakes, as well as regions associated with seasonal wetlands (Prigent et al. 2007). This class is used to represent large areas that are inclined to have standing water, in particular in the North American continent. Figure 2 shows the results of the climatological classification for June at a $0.25^\circ \times 0.25^\circ$ equal-area spatial resolution.

The surface class classification is used as a climatology. This means that the surface type for a satellite field of view is only determined by its location and month. This is already an improvement compared to classical surface type classifications that are constant in time. A potential improvement would be to estimate the surface class based on real-time emissivities, but this would be complicated since the estimation of these emissivities is difficult for cloudy situations; although this is possible (Aires et al. 2001), the uncertainties are higher.

e. Sensitivity of the microwave observations to the clouds

To assess the sensitivity of each microwave channel to the presence of clouds, histograms of the brightness temperatures (Tb's) are examined, for the four SEVIRI-derived cloud classes. Figure 3 shows a representative subset of these histograms, for land and ocean separately, for the whole data period June–October 2006. In the most transparent channels (23.8, 50.3, and 89 GHz), the surface contribution to the signal is obvious, with very different ranges of temperatures over land and ocean, along with significantly less sensitivity to the cloud presence and types over land (note that for a given frequency channel, the histograms are drawn on the same scales for an easier comparison between land and ocean cases). Over a high-emissivity surface such as land (i.e., a radiometrically warm surface), the presence of clouds tends to decrease the observed Tb's even at low frequencies where scattering does not play a significant role (up to 80 GHz). The opposite occurs over the radiometrically cold ocean, where clouds generally induce an increase in the Tb's, especially for the horizontal polarization. At 89 GHz, scattering by the cloud ice phase starts to interfere significantly with the signal, and this is especially visible over land where the decrease of the Tb's for high clouds contrasts with the warm background. This property will make high clouds easier to identify over land than over ocean with frequencies around 89 GHz. For more opaque channels, the difference in the behavior between land and ocean is less clear and decreases with channel opacity. Note that over both land and ocean, the histograms at 150 GHz and above are rather well separated for the medium and high cloud types, making these channels likely very promising for the classification of these cloud types.

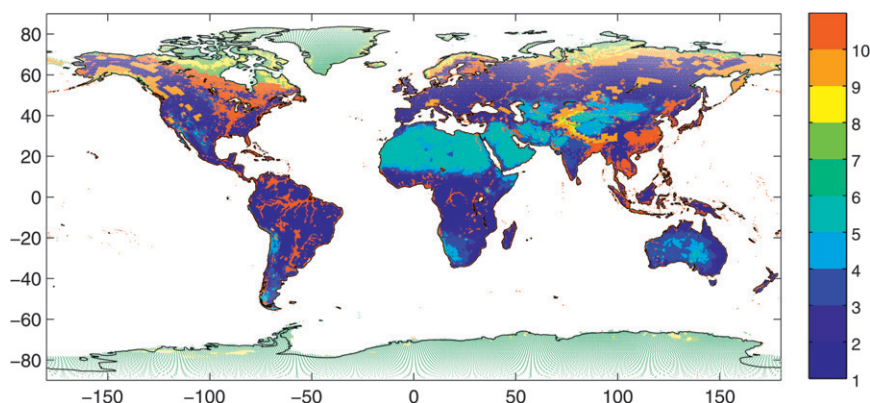


FIG. 2. Surface emissivity classification for June. Surfaces 1–5 correspond to decreasing vegetation densities, from dense forest to desert. Surfaces 6–9 are for snow- and ice-covered pixels. Surface 10 corresponds to pixels that are partly covered by standing water.

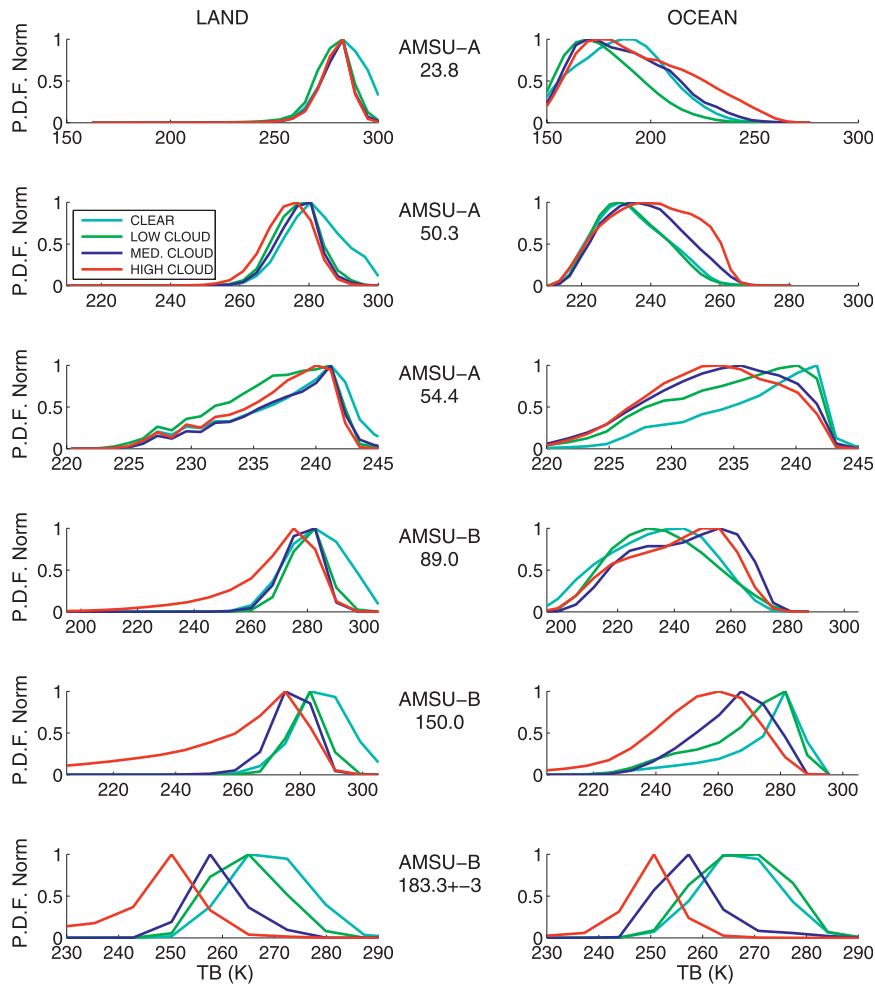


FIG. 3. Brightness temperatures histograms for the 4 SEVIRI-derived cloud classes, over (left) land and (right) ocean for window channels 23.8, 50.3, 89, and 150 GHz; and for sounding channels at 54.4 and 183.3 ± 3 GHz.

This analysis shows that there is potential for cloud classification with these microwave observations, but that one frequency channel alone will not provide all the information. As a consequence, we suggest a multispectral analysis. All AMSU-A and -B information will be used simultaneously. At the end of this study (section 4e), the potential of different channel groups will also be tested.

3. Classification methods

Let x be a vector of p features (i.e., the microwave observations in this case). A dataset $\mathcal{B} = \{(x^i, y^i); i = 1, \dots, M\}$ of M couples relates these observations x^i to labels y^i where $y \in \{1, \dots, N\}$ represents N groups (i.e., the cloud classes). These N groups are defined a priori. A supervised classification algorithm is a statistical procedure that allocates any vector of observations x onto the N groups.

The training of the classifier is based on the use of the “training” dataset \mathcal{B} of previously labeled items. The goal of the classifier is to reproduce, as well as possible, the a priori classification in the training dataset and also to extrapolate this classification to new observations, outside the training dataset.

In this application, the groups are the $N = 4$ cloud classes ($y = 1$ for clear, 2 for low, 3 for medium, and 4 for high clouds). The a priori classification y is provided by the transformed SEVIRI classification (section 2a) and the observations x are the microwave observations from AMSU-A and -B described in section 2b.

a. Linear and quadratic discriminant analyses

There are two approaches to defining classification procedures: one based on statistics, in particular Bayesian statistics, and the other one based on geometric considerations. Both are linked since each statistical hypothesis

(e.g., equality of intraclass variance, Gaussian character of the intraclass variance) uses geometrical considerations and some statistical optimization.

The linear discriminant analysis (and the Fisher's linear discriminant) is a statistical method used to find the linear combination of the P features that best separates the N groups. The resulting combination defines P -hyperplane "frontiers" that separate the points from each group. In the case where $P = 2$, the frontiers are lines, and in the case where $P = 3$, they are planes.

For the quadratic classifier, the frontiers are assumed to be quadratic in the space of the observations. In the special case where the observations have two features (i.e., $P = 2$), the surfaces separating the classes are conic sections (i.e., a line, a circle or ellipse, a parabola, or a hyperbola).

The representation the capacities of the linear classifier are included in the quadratic classifier (lines can be special cases of conic sections). It means that the quadratic approach should give equal or better results than the linear discriminant analysis.

Another important feature to be mentioned is that these linear and quadratic classifiers also provide the a posteriori probability for an observation x to be in each one of the M classes. This will be exploited in the following. It is not the goal of this paper to present thoroughly these classical methods; the interested reader can refer to (Krzanowski 1988; Seber 1984).

b. Neural network method

A neural network (NN) is a nonlinear mapping model. The multilayered perceptron (MLP) is adopted here (Rumelhart et al. 1986). The inputs x are the predictors and they represent any source of information for the prediction; the outputs y represent the predicted variables. In our case, x is composed of microwave observations from the AMSU-A and -B instruments. In the following, the number of inputs will vary, depending on the available observations: 20 for AMSU-A and -B, 15 for AMSU-A only, 5 for AMSU-B only, and 6 for the window channels (see section 4e). This allows us to test separately the contribution of each instrument. Each NN has a unique hidden layer with 20 neurons. The prediction y in the NN output is a four-digit binary code that represents the cloud classes. This coding in the NN output is the best way to obtain a classifier able to estimate a posteriori probabilities (Bishop 1996). In this way, and similarly to the linear and quadratic classifiers, the NN classifier is able to provide a predicted class y for each new observation x , along with the a posteriori probability for x to be in each one of the $N = 4$ cloud classes. This is a very interesting feature that can be used to refine the classification results.

The NN is trained to reproduce the behavior described by the learning dataset \mathcal{B} . Provided that enough

samples (x^i, y^i) are available, any continuous relationship as complex as it is can be represented by an MLP (Hornik et al. 1989; Cybenko 1989). The initial weights of the NNs are randomly chosen using a uniform distribution. The method used to perform this training is the classical back-propagation algorithm (Rumelhart et al. 1986). It is an optimization procedure that is perfectly adapted to the MLP architecture. It is designed to minimize a loss function (i.e., quality criterion), the least squares differences between the desired and the NN estimated outputs. The stopping criterion used to terminate the learning is based on the gradient of the loss function. The learning procedure uses the following components:

- a learning dataset to estimate the parameters of the NN (representing 60% of the original satellite dataset);
- a generalization dataset (20% of the original dataset) to estimate the ability of the NN to classify data that are not present in the learning dataset; the generalization errors are used to identify the best NN architecture; and
- a validation dataset (again, representing 20% of the original dataset) used to obtain reliable error estimates; in section 4, all of the presented statistics will be estimated with this validation dataset.

The results presented are estimated on the validation dataset.

c. Uniformization of the learning dataset

A classical difficulty in statistical techniques appears also in the classification problem. If regimes or classes are underrepresented in the dataset used to train the classifier, they have a limited incidence on the classifier design and performance. If the goal of the statistical retrieval is to perform uniformly well for each of the N classes, the number of points in each class needs to be uniformized in the learning dataset. Otherwise, the highly populated classes will have too much weight and will drive the classifier. It will predict these classes too often and underestimate the occurrence of the other classes. In this study, all the of classification training will be performed with uniformized learning datasets.

4. A cloud classification and mask from MW observations

In the following sections, the results will be given using the validation dataset (i.e., data that are not used during the learning stage or during the NN architecture optimization). Furthermore, results will be presented on uniform and nonuniform pixels.

a. Choosing the classifier model

The three classification methods (section 3) are tested to reproduce the SEVIRI-derived cloud classification, using the microwave observations. Since the surface contribution is large for some of the microwave observations, land and sea are separated to facilitate the classification. The cloud classification over land is first trained and tested on the entire land dataset. Results are analyzed over not only all land pixels but also for individual surface types, as described in section 2d.

The available datasets of section 2 are divided into a “learning set” to train the classifiers, a “testing set,” and a “validation set” to test the classifiers (see section 3b). Each day of the dataset is randomly chosen to be part of the learning or the testing set. In this way, the testing of the algorithms can be undertaken in a particular day and a map of the results can be obtained. The division of the three datasets is random for each day.

The confusion matrix is a measure of the classification error statistics. For an original cloud class, it provides the percentage of well-classified pixels, together with the repartition of the misclassified points in the other classes. A perfect classifier has a diagonal confusion matrix with 100% well-classified statistics, for the clear sky and low-, medium- and high-cloud classes. The statistics are performed on the datasets described in section 2; that is, the data have been filtered to take into account only the homogeneous clear-sky and high, middle, and low opaque cloud scenes.

Table 4 compares the confusion matrix of the three classifiers, trained globally over land and ocean separately: linear, quadratic, and NN “global.” Snow classes are suppressed from the analysis. Snow has a very variable microwave emissivity that is difficult to handle in atmospheric retrievals from microwave observations. In addition, the use of a geostationary satellite as a reference for the cloud classification is not ideal for snowy regions, due to its limited coverage of the high latitudes. Polar orbiters should be preferred for these specific cases. Furthermore, some difficulties occur in some regions, for example, in the central United States and eastern China where snow pixels are indicated to be present. These misclassified pixels are related to problems in the NSIDC snow flag that is used in our classification (section 2d). These difficulties should be solved in our next version of the surface-type classifications where only the microwave emissivities will be used. This work is a first step, but snow-covered regions could be the subject of a future study.

Regardless of the classification method, clear situations are easier to identify over ocean than over land, due to the lack of contrast between the atmospheric and surface contributions over land. Low-level clouds are

generally more difficult to detect than the higher clouds. The quadratic classifier performs better than the linear classifier. The quadratic classifier allows for more complex class frontiers than the linear option. Since linear classification is a subcase of the quadratic model, linear results should always be of lower quality than quadratic ones. The NN classification outperforms both linear and quadratic classifiers. It is able to represent quadratic separators so its results should be at least comparable to the quadratic results. In addition, the highly nonlinear nature of the NN makes it possible to be “situation dependent”: The classifier structure is not the same for very different conditions. This means that the NN can more easily adapt itself to different surface conditions. With the linear and the quadratic models over land, the results can depend on the surface type (e.g., less than 10% of well-classified low-level clouds over deserts and more than 50% over dense forests). With the NN classifiers, the results are more even for the different land surface types.

Since the NN appears to be the best method, the classification experiments in the following sections will be performed using only the NN approach.

It can be noted that even with the NN (global), low clouds are difficult to identify correctly. This is particularly true over desert (surface type 5), where only 35% of the initial pixels are correctly classified. Overall, there is considerable confusion between the clear and the low-cloud cases. Over a strongly emitting surface, low clouds that emit at a rather similar temperature do not show any significant contrast. Sounding channels with weighting functions peaking rather low are also affected by the surface contribution and have difficulties when isolating the low-cloud signal.

To reduce this difficulty, a specialized NN is developed for each land surface type. A classifier is defined for each surface type of section 2d. Each observation is processed either by the general classifier, or by its specialized classifier (its class is defined using the surface-type class climatology, so it is determined based on the location and month of the year). The results of this new set of classifiers are represented in the last column of Table 4 (labeled NN by land surf). The rate of well-classified pixels for the low-cloud class over surface type 5 is increased from 35% to 69%. Surprisingly, some results are slightly degraded when the NN classifier was specialized. This degradation can come for various reasons. One of the reasons is that the classifier is now dependent on the surface-type classification, which can be misleading, although the surface information (section 2d) is dynamical and depends on the month of the year, thus, limiting this problem. The number of points available for the training of the classifier is reduced, which can be negative for surface classes that have limited populations

TABLE 4. Cloud classification confusion matrices for the linear, quadratic and NN global classification, as well as NN by surface class, over land and ocean. Surfaces 1–5 correspond to decreasing vegetation densities, from dense forest to desert, and surface 10 corresponds to pixels partly covered by standing water. Snow- and ice-covered pixels are omitted. These statistics are performed on relatively homogeneous pixels (see section 2c and Table 3). Boldface signify the diagonal line of the confusion matrices.

		%	Linear global				Quadratic global				NN global				NN by land surface				No. pixels
			Clear	Low	Medium	High	Clear	Low	Medium	High	Clear	Low	Medium	High	Clear	Low	Medium	High	
Land	Surface 1	Clear	37.07	25.86	36.71	0.36	58.24	22.66	18.46	0.63	70.58	22.78	5.54	1.10	81.22	14.79	2.96	1.04	341 408
		Low	9.29	53.02	37.55	0.15	24.09	55.06	20.57	0.27	11.60	74.04	13.79	0.56	14.44	71.70	13.38	0.48	40 636
		Medium	3.17	24.66	70.06	2.11	6.27	14.32	76.97	2.44	2.67	14.28	76.82	6.23	3.07	11.56	80.01	5.35	19 759
		High	0.49	6.80	24.65	68.05	1.33	5.39	26.16	67.12	1.46	2.84	7.70	87.99	2.28	2.26	7.92	87.54	44 398
Surface 2	Clear	54.30	25.80	19.48	0.42	66.68	24.57	8.13	0.62	76.43	19.68	3.36	0.53	81.29	14.50	3.22	0.99	220 080	
	Low	9.32	68.47	21.89	0.32	15.94	71.41	12.18	0.47	10.52	79.46	9.41	0.60	12.37	75.38	10.90	1.35	23 168	
	Medium	4.86	31.59	61.08	2.47	6.75	19.92	70.23	3.10	3.18	16.76	74.12	5.94	3.77	11.72	76.99	7.52	10 342	
	High	1.98	15.51	21.63	60.88	3.23	12.52	28.94	55.31	3.29	5.27	9.45	81.99	3.53	2.99	7.37	86.10	16 300	
Surface 3	Clear	69.33	14.58	15.69	0.40	81.04	11.98	6.22	0.76	85.35	11.27	3.11	0.27	84.05	12.33	3.33	0.28	179 880	
	Low	11.14	62.62	25.19	1.04	16.12	67.96	14.81	1.11	14.79	75.64	9.22	0.35	8.77	82.04	8.57	0.61	5402	
	Medium	8.14	15.69	72.72	3.45	4.86	11.10	81.11	2.93	3.13	10.02	82.01	4.83	3.13	9.24	82.10	5.53	4469	
	High	5.82	5.25	20.19	68.75	6.96	6.06	24.60	62.38	6.38	2.77	8.64	82.21	6.50	2.44	7.47	83.59	7257	
Surface 4	Clear	77.77	10.94	11.02	0.28	87.83	7.31	4.37	0.49	86.69	10.90	2.23	0.19	86.56	10.36	2.79	0.29	129 397	
	Low	23.64	50.32	25.92	0.13	26.04	54.61	18.96	0.38	21.87	62.71	15.17	0.25	16.69	72.31	10.75	0.25	791	
	Medium	15.03	7.01	72.98	4.98	4.01	3.00	89.17	3.82	1.79	3.00	90.24	4.98	1.50	4.20	89.95	4.35	2069	
	High	5.24	1.53	16.68	76.54	6.56	1.32	22.71	69.41	4.49	1.07	6.45	87.99	3.78	1.03	6.35	88.84	2805	
Surface 5	Clear	88.79	7.37	3.70	0.15	95.14	2.95	1.74	0.17	88.23	10.16	1.49	0.11	80.93	16.21	2.76	0.10	507 674	
	Low	77.21	9.40	12.37	1.02	76.56	9.95	13.12	0.37	50.23	35.72	12.37	1.67	17.58	69.21	11.91	1.30	1075	
	Medium	28.49	5.90	57.97	7.64	7.70	0.64	87.47	4.19	2.85	2.30	88.30	6.54	2.30	2.331	89.88	5.49	4556	
	High	4.49	1.34	14.83	79.35	5.44	1.22	27.82	65.52	2.05	1.00	5.68	91.27	1.98	0.83	6.27	90.93	4101	
Surface 10	Clear	45.85	28.45	24.11	1.59	57.22	29.53	11.22	2.03	71.11	23.73	4.31	0.84	80.04	14.62	4.56	0.78	9198	
	Low	8.03	62.20	28.13	1.64	11.46	70.68	16.21	1.65	11.22	76.72	11.27	0.79	12.81	73.49	12.87	0.84	8581	
	Medium	3.81	30.89	61.76	3.53	6.61	19.94	69.23	4.22	3.40	18.01	71.77	6.81	3.48	11.86	78.30	6.35	3936	
	High	0.92	9.19	21.72	68.17	2.43	8.87	22.31	66.39	2.47	3.91	7.50	86.11	3.10	3.13	7.73	86.04	9427	
Total land	Clear	65.48	17.02	17.12	0.37	77.42	14.03	7.96	0.59	80.75	15.59	3.18	0.48	81.87	14.53	3.06	0.53	1 474 280	
	Low	10.32	58.61	30.64	0.43	20.34	61.86	17.23	0.58	12.09	75.26	12.04	0.60	13.31	73.67	12.23	0.79	80 606	
	Medium	7.11	23.29	66.45	3.15	6.21	13.98	76.70	3.11	2.87	13.02	78.02	6.09	3.15	10.09	80.71	6.05	45 998	
	High	1.63	8.20	22.75	67.41	2.66	6.91	26.26	64.17	2.48	3.28	8.07	86.17	3.02	2.42	7.62	86.94	85 022	
Ocean	Total ocean	Clear	87.34	8.06	2.51	2.08	88.41	6.95	2.82	1.81	87.99	7.46	2.61	1.94	724 979
		Low	21.08	55.92	21.33	1.67	21.70	63.56	13.06	1.68	9.93	79.91	9.54	0.62	528 995
		Medium	7.31	8.82	71.94	11.92	9.71	12.88	69.40	8.01	1.70	9.48	77.56	11.26	47 128
		High	4.20	0.56	19.91	75.32	7.53	3.97	25.20	63.31	2.12	2.69	11.13	84.07	114 020

(the generalization errors are always compared to the learning errors in order to check if the number of samples is enough to allow the classifier to learn). The training of the NN using all points from all of the surface types can also be more adequate when learning global sensitivities that can become too subtle when looking only to one surface type. However, the overall effect of the specialization of NNs by surface type is largely positive¹ and is adopted in the following sections.

b. Reducing uncertainty by using the a posteriori probabilities

The classification techniques tested in this study, in particular the selected NN approach, can provide estimations of the a posteriori probability of one observation being in each of the N possible classes: $P(y = c_i/x)$ for $i = 1, \dots, N$. The class chosen by the classifier is the most probable class, but these probabilities can also be used to control the uncertainty of the classification. When two or more probabilities are comparable, the situation is ambiguous. To reduce classification ambiguities, one can use the classification only when the best a posteriori probability is higher than a predetermined threshold (between $1/N$ and 1). This decrease in uncertainties improves the accuracy of the classification but the number of classified points is reduced because they are too ambiguous.

In Table 5, the confusion matrices are given when no a posteriori probability filtering is used, and with a posteriori thresholds of $h = 0.4$ and 0.8 . Statistics remain stable when the threshold is lower than or equal to 0.4 . This means that the classifier usually is able to retrieve a class with more than 0.4 a posteriori probability. The quality of the classification statistics improves significantly when a threshold larger than $h = 0.4$ is chosen. The percentage of well-classified situations (on the diagonal of the confusion matrices) increases by more than 10% when using a threshold $h = 0.8$. From Table 5, it is possible to state that the classifier can separate quite precisely the ambiguous situations from the highly probable ones. Table 5 also provides the number of classified situations, after the a posteriori probability filtering.

Figure 4 represents, on the left side, the evolution of the number of situations kept after the a posteriori filtering, when the threshold h increases, for clear pixels and low, medium, and high clouds (A, B, C, and D). As commented upon previously, this number stays relatively constant for $h \leq 0.4$. For higher thresholds, the decrease

is quite linear, which simplifies the choice of the adequate h . There are some differences between the surface types. For example, the high-cloud classification is more ambiguous over ocean than over land, which is contrary to the case of clear pixels, which are more ambiguous over land than ocean. Figure 4 also shows on the right side the evolution of the percentage of well-classified pixels with an increasing threshold h : For clear and low-, medium-, and high-cloud pixels (E, F, G, and H). The spread is more important for low thresholds of h and converges to higher values when h increases. There is a compromise to be found between a more precise classification with good rates of well-classified pixels (larger h) and a large number of pixels being treated by the classifier (smaller h). No general rule can be set, and the right compromise depends on the application for which the classification is being used. In the following, a threshold of $h = 0.4$ is adopted in order to process enough pixels to produce classification maps for the analysis.

c. Analysis of the resulting classification maps

In Fig. 5, the cloud classifications are tested using coincident SEVIRI and microwave (i.e., AMSU-A and AMSU-B) observations on 1 July 2006. This day is not included in the learning dataset used to train the cloud classification algorithms (see section 3b). The SEVIRI-derived four-class classification scheme is shown in Fig. 5a, and the MW-retrieved classification is represented in Fig. 5b. Note that the first map is not an “instantaneous” map from the SEVIRI observations. It shows the SEVIRI classification only for the pixels with a space-time coincidence with the MW observations during the whole day of 1 July 2006. The maps represent the type of data that have been included in the learning dataset used to train the microwave classifier. In particular, the homogeneity tests in section 2c have been used. Since SEVIRI has a repeat cycle of 15 min, the temporal coincidence with MW observations is kept under 7.5 min. Some minor differences can be observed with, for example, low clouds misclassified by the microwave over the desert Arabian Peninsula. However, the overall agreement is good and confirms the good confusion matrix statistics presented in section 4a.

To investigate the behavior of the MW classifier in heterogeneous situations (i.e., various cloud types in an MW field of view), some statistics are performed on coincident SEVIRI–AMSU observations without the filtering presented in section 3c. Figure 6 investigates the SEVIRI cloud classification at 1230 UTC 1 July 2006 (left) together with the AMSU classification (right) for an orbit that passes over Madagascar at the same time. The temporal coincidence is not as good as in Fig. 5 where coincidences were kept under 7.5 min. Some

¹ This specialization of the land classifiers would be even more beneficial for the linear and quadratic classifiers (not shown).

TABLE 5. Cloud classification confusion matrices for the neural network by surface class over land and ocean with no a posteriori threshold used with a posteriori = 0.4 and 0.8. These statistics are performed on relatively homogeneous pixels (see section 2c). Boldface signify the diagonal line of the confusion matrices.

		No a posteriori				A posteriori = 0.4				A posteriori = 0.8						
		Clear	Low	Medium	High	No. pixels	Clear	Low	Medium	High	No. pixels	Clear	Low	Medium	High	No. pixels
Surface 1	Clear	81.22	14.79	2.96	1.04	341 408	82.15	14.34	2.54	0.98	334 427	95.78	2.74	0.53	0.95	150 784
	Low	14.44	71.70	13.38	0.48	40 636	14.29	72.28	13.06	0.38	39 794	5.53	85.69	8.69	0.08	12 251
	Medium	3.07	11.56	80.01	5.35	19 759	2.71	11.04	80.92	5.33	19 318	0.59	2.72	94.16	2.53	9452
Surface 2	High	2.28	2.26	7.92	87.54	44 398	2.11	1.89	7.65	88.34	43 855	0.72	0.23	2.05	97.00	35 585
	Clear	81.29	14.50	3.22	0.99	220 080	81.96	14.18	2.94	0.92	216 700	96.84	2.36	0.44	0.35	105 125
	Low	12.37	75.38	10.90	1.35	23 168	12.29	75.85	10.62	1.25	22 783	89.59	2.04	4.97	0.17	6666
Surface 3	Medium	3.77	11.72	76.99	7.52	10 342	3.46	11.40	77.68	7.46	10 167	0.57	2.04	93.38	4.00	4697
	High	3.53	2.99	7.37	86.10	16 300	3.41	2.65	7.21	86.73	16 124	1.78	0.26	2.01	95.95	11 978
	Clear	84.05	12.33	3.33	0.28	179 880	85.05	11.84	2.86	0.26	174 917	97.13	2.38	0.31	0.18	92 588
Surface 4	Low	8.77	82.04	8.57	0.61	5402	8.41	82.85	8.19	0.55	5289	3.69	92.57	3.60	0.13	2276
	Medium	3.13	9.24	82.10	5.53	4469	2.79	8.60	83.07	5.54	4372	0.21	2.37	95.10	2.32	2367
	High	6.50	2.44	7.47	83.59	7257	6.06	2.13	7.23	84.59	7151	2.81	0.36	1.81	95.02	5583
Surface 5	Clear	86.56	10.36	2.79	0.29	129 397	87.39	9.79	2.60	0.23	126 573	96.40	2.85	0.57	0.17	68 592
	Low	16.69	72.31	10.75	0.25	791	16.54	72.39	10.81	0.25	786	4.62	90.02	5.35	0.00	411
	Medium	1.50	4.20	89.95	4.35	2069	1.23	3.49	91.11	4.17	2037	0.07	0.57	97.92	1.44	1392
Surface 10	High	3.78	1.03	6.35	88.84	2805	3.01	0.62	6.32	90.05	2753	0.99	0.13	1.98	96.90	2319
	Clear	80.93	16.21	2.76	0.10	507 674	81.20	16.14	2.57	0.09	504 005	93.24	5.95	0.73	0.07	186 896
	Low	17.58	69.21	11.91	1.30	1075	17.64	69.51	11.54	1.31	1066	4.57	84.41	11.02	0.00	372
Surface 10	Medium	2.30	2.33	89.88	5.49	4556	1.89	2.00	90.61	5.50	4507	0.03	0.10	97.52	2.35	3022
	High	1.98	0.83	6.27	90.93	4101	1.81	0.74	6.13	91.32	4079	0.24	0.15	1.83	97.78	3379
	Clear	80.04	14.62	4.56	0.78	91 982	81.31	13.93	4.01	0.75	89 304	96.25	2.74	0.47	0.54	42 599
Surface 10	Low	12.81	73.49	12.87	0.84	8581	12.66	74.00	12.57	0.77	8358	7.16	88.10	4.49	0.25	2403
	Medium	3.48	11.86	78.30	6.35	3936	2.80	11.34	79.50	6.36	3819	0.73	1.73	94.01	3.53	1503
	High	3.10	3.13	7.73	86.04	9427	3.00	2.77	7.46	86.77	9312	1.05	0.35	1.51	97.09	7211
Total land	Clear	81.87	14.53	3.06	0.53	1 474 280	82.54	14.21	2.75	0.50	1 449 708	95.51	3.57	0.54	0.38	648 093
	Low	13.31	73.67	12.23	0.79	80 606	13.17	74.22	11.91	0.70	79 007	5.40	87.77	6.71	0.12	24 621
	Medium	3.15	10.09	80.71	6.05	45 998	2.78	9.60	81.61	6.01	45 066	0.45	1.97	94.68	2.90	22 742
Total ocean	High	3.02	2.42	7.62	86.94	85 022	2.82	2.07	7.39	87.72	83 995	1.10	0.25	1.94	96.70	66 516
	Clear	87.99	7.46	2.61	1.94	724 979	88.62	7.21	2.34	1.83	716 432	97.45	1.33	0.27	0.95	413 328
	Low	9.93	79.91	9.54	0.62	528 995	9.84	80.24	9.36	0.55	524 493	5.25	92.27	2.39	0.09	231 152
Total ocean	Medium	1.70	9.48	77.56	11.26	47 128	1.58	9.33	77.87	11.22	46 714	0.89	4.17	89.44	5.50	19 040
	High	2.12	2.69	11.13	84.07	114 020	1.96	2.47	10.90	84.67	112 743	0.72	0.35	2.08	96.85	74 338

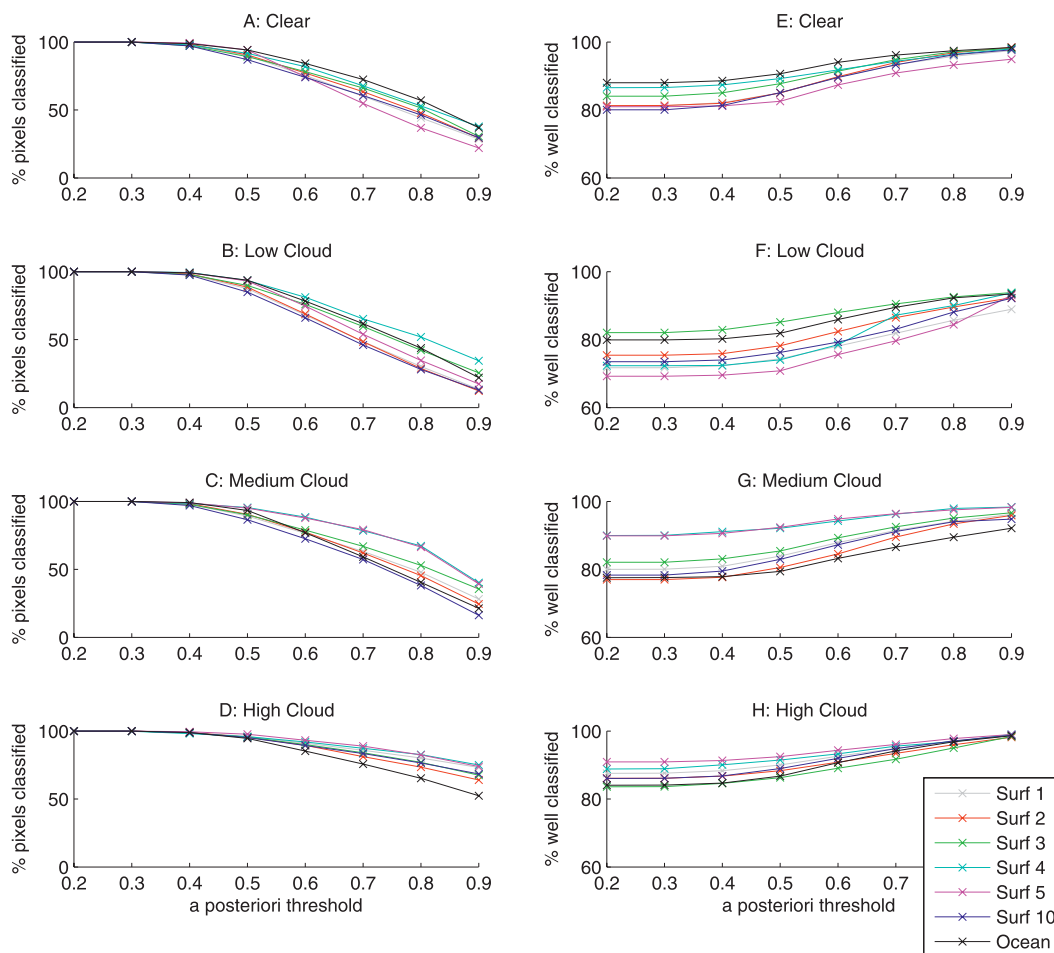


FIG. 4. (left) Number of classified pixels and (right) percentage of well-classified pixels after a posteriori probability filtering with an increased threshold. (top to bottom) Statistics are for clear sky and low, medium, and high clouds for each surface classes (i.e., vegetated, 1–5; with standing water, 10) and for ocean surfaces.

misclassified pixels can be observed in this figure, such as the presence of low clouds in the Arabic Peninsula already mentioned. Another difference is observed in Africa, west of Madagascar, where medium-altitude clouds do not appear in the microwave classification. It can be noted in Fig. 5 that no point was located in this region. It means that these points have been filtered out by the quality criteria of section 2c; that is, the cloud classification was too heterogeneous in space. Since the SEVIRI ensemble of cloud classes inside the AMSU pixels is ambiguous, the microwave classifier has to integrate its response spatially. Therefore, low clouds are found instead of clear sky and medium-altitude clouds. This problem results from the differences in the horizontal resolutions of the different instruments and cannot be solved. It should be noted however that even if middle clouds are misclassified as low, this still represents a successful cloud detection. This can be of great

use; forthcoming retrieval algorithms may use the classification to skip these pixels (cloud mask will be presented in the following section).

In the remainder of this section, the comparative analysis of the MW and SEVIRI cloud classifications will consider AMSU classification pixels with an a posteriori threshold quality criteria fixed to 0.4 (see section 4b). The SEVIRI class associated with the AMSU pixel is the more frequent class among the SEVIRI pixels falling into the AMSU pixel. If the frequency of this class fulfills the criteria given in section 2c, the AMSU pixel is labeled homogeneous otherwise it is labeled heterogeneous.

Figure 7 gives, for each SEVIRI class over ocean, the distribution of the MW classes, for the homogeneous (panel a) and the heterogeneous (panel b) cases. As expected, over ocean and over land (figure not shown), the statistics for homogenous cases are in agreement

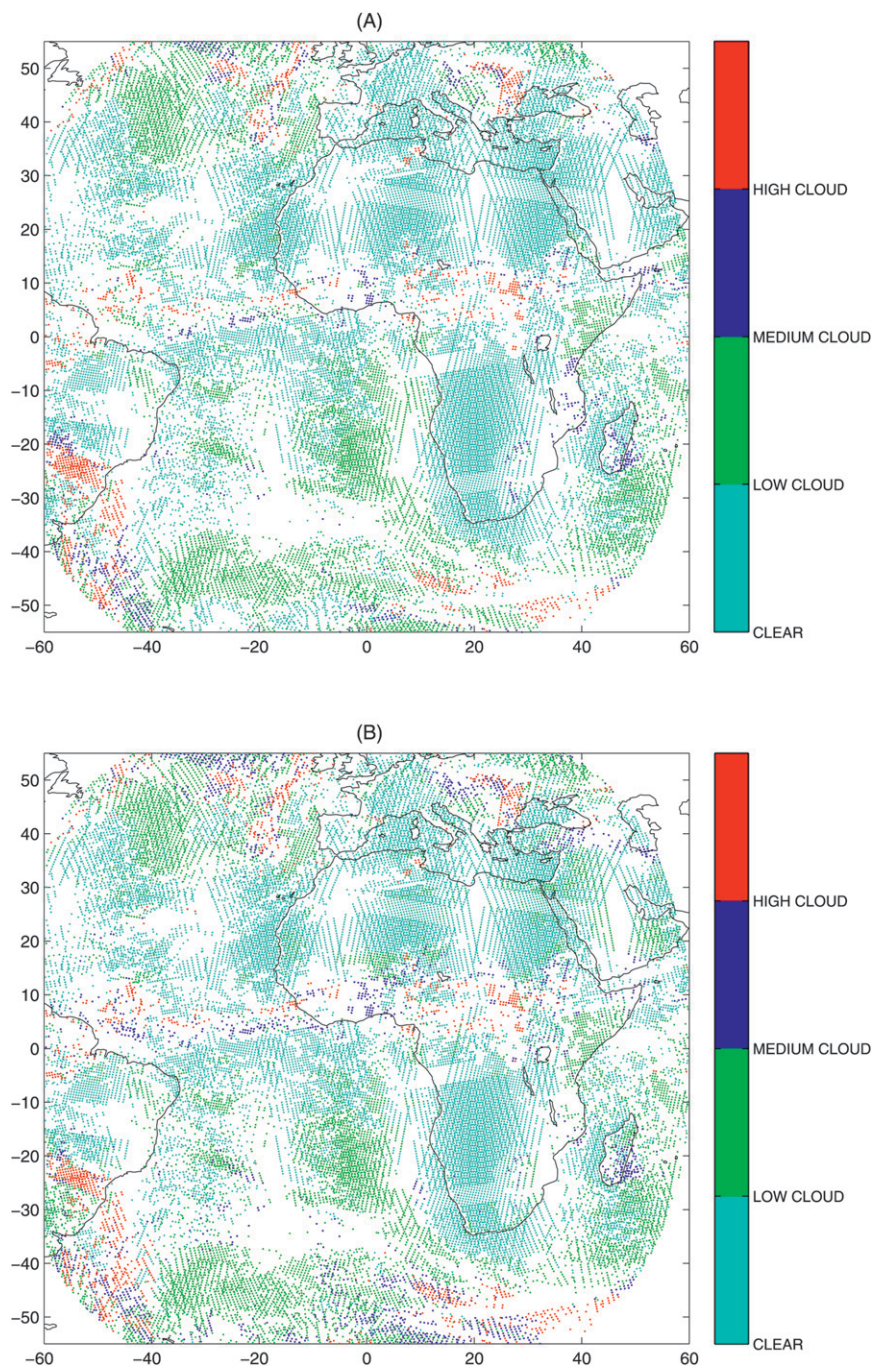


FIG. 5. (a) Classification from SEVIRI and (b) retrieved classification from microwave observations for 1 Jul 2006.

with the values given in Table 5 considering the 0.4 a posteriori quality test value. Not surprisingly, for heterogeneous cases, the agreement between the MW and SEVIRI classes decreases. However, the well-classified probabilities remain above 50%. It is important to understand here that this result includes the errors of the MW classifier; in addition, it includes the spatial dispersion of

the SEVIRI classification inside a bigger MW pixel. Figure 7 shows that, for both homogeneous and heterogeneous cases, partially cloud-covered pixels are mostly classified as clear by the AMSU classifier. For thick cirrus, the most frequent AMSU cloud class is high cloud. However, the thin cirrus clouds are classified as clear by the MW classifier in more than 70% of the cases, in agreement

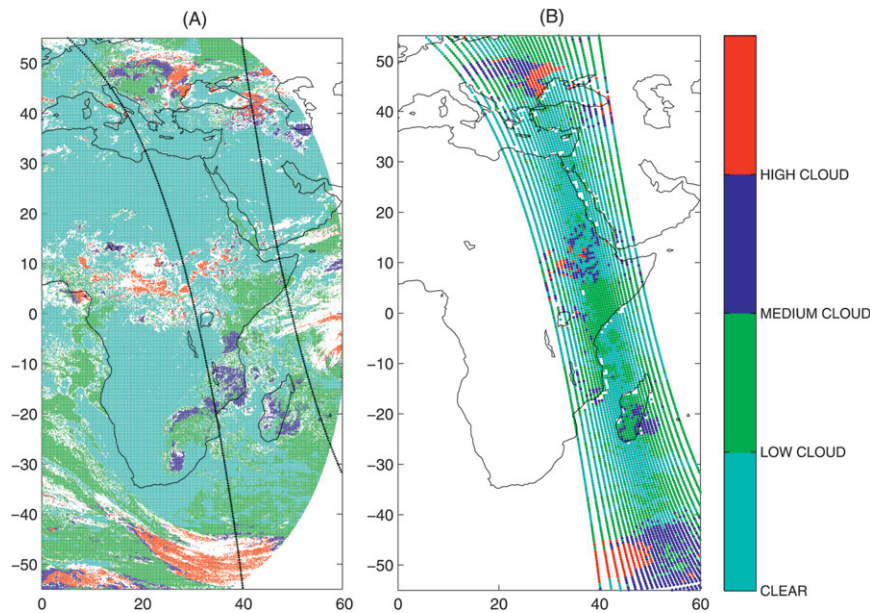


FIG. 6. (a) Classification from SEVIRI and (b) retrieved classification from AMSU-A and -B over a NOAA-16 orbit at 1230 UTC 1 Jul 2006.

with the fact that cirrus clouds are transparent for microwave measurements when their emissivity is small.

Using the results of Fig. 7, a new SEVIRI classification has been defined to characterize each of the MW classes (clear, low, middle, and high), the associated SEVIRI cloud type, or mixing of cloud types. The thin cirrus and partially covered classes have been reclassified as clear and the thick cirrus have been kept as high opaque cloud. Figure 8 shows the distribution of these new SEVIRI classes in each MW cloud label. Clear sky from the MW classifier corresponds to SEVIRI clear sky, partially covered pixels or thin cirrus in more than 80% of the cases. High clouds are associated with SEVIRI high clouds and/or thick cirrus in more than 70% of the cases. For MW low- and midlevel clouds, there is a difference in behavior between land and ocean. The MW classifier detects relatively well the low clouds over ocean. This is not the case over land: the low- and midlevel cloud MW classifications are frequently covered by SEVIRI clear and/or broken clouds and/or thin cirrus. It can also be noted that the frequencies of the clear-sky and low-, middle-, and high-cloud MW classes over ocean are in relatively good agreement with the equivalent distribution in ISCCP stage DX data over the same region (Sèze et al. 2009). Low clouds are the more frequent cloud type over ocean. Over land, the frequent presence of thin cirrus strongly biases the MW high-cloud frequency toward low values compared to the high-cloud frequencies observed with a visible and an infrared radiometer.

Once the MW cloud classification is trained over the SEVIRI disk observation, it is possible to extend it to the global coverage of the NOAA-16 platform. If the classifier has learned the correct relationships between the MW observations and the cloud classes, then it should be possible to use these same relationships in different geographical locations. Figure 9 represents such an extension for the same day, 1 July 2006. The multiple orbits are not coincident in time and there are overlapping orbits, especially at high latitudes. The spatial structures of the cloud classes are consistent and show overall realistic continuity at the transition between land and ocean. Other locations are well retrieved by the microwave classification, even if the time differences become greater. It is expected that the quality of the classification is degraded at high latitudes ($>50^\circ$) as it has been trained over the SEVIRI observation disk, between -50° and $+50^\circ$ in latitude, mostly over Africa. To extend the cloud classification model to higher latitudes, snow- and ice-covered pixels should be included in the dataset \mathcal{B} .

d. Cloud mask

The classifier defined in this study retrieves four cloud classes: clear sky and low, medium, and high clouds. This classifier can also be used as a cloud mask. Two schemes are compared in this study: 1) the situation is labeled as clear if the a posteriori probability (see section 3b) to be clear is the highest and 2) the a posteriori probabilities of being low, medium, or high cloud are summed and if this

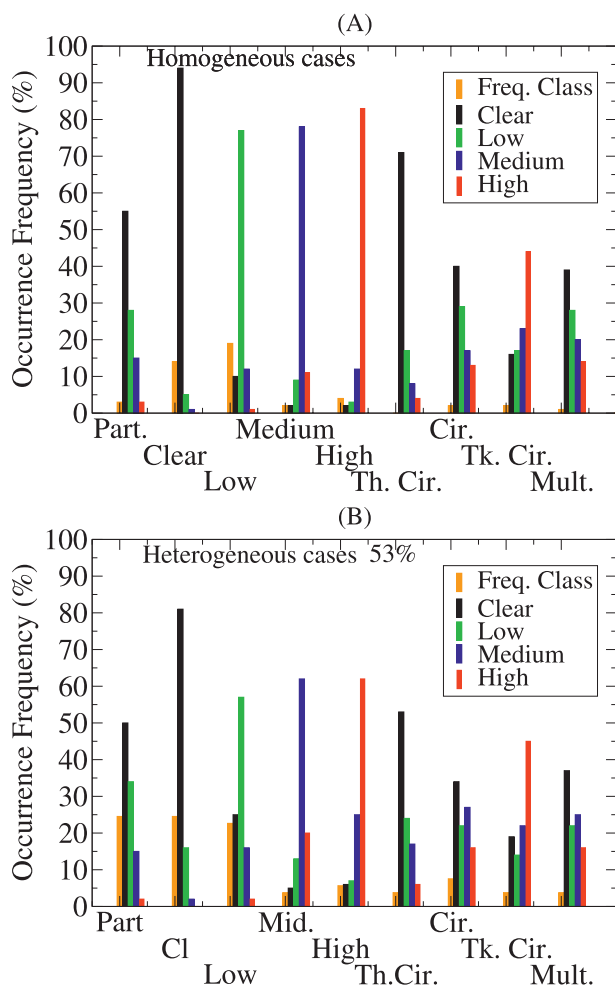


FIG. 7. Occurrence frequency vs cloud class for (a) homogeneous and (b) heterogeneous AMSU pixels for the MW cloud class distribution for each SEVIRI cloud class over ocean. The frequency of this class in the AMSU pixel set (orange bar) is also indicated.

total is higher than 0.5, then the situation x is classified as cloudy (or clear in the opposite situation). It is expected that the second scheme that sums the a posteriori probabilities of being a low, medium, or high cloud retrieves more cloudy situations than the first scheme.

Similar to Fig. 5 where the cloud classifications were compared, Fig. 10 represents the cloud masks using

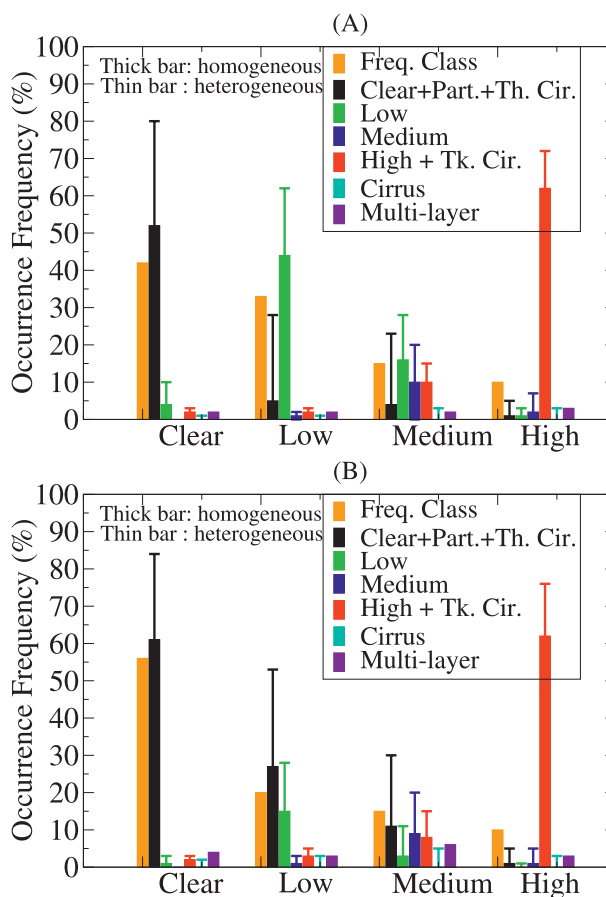


FIG. 8. SEVIRI cloud class frequency distribution for each MW cloud class, over (a) ocean and (b) land. The frequency of this class in the AMSU pixel set is also provided (orange bar).

coincident SEVIRI–MW observations on 1 July 2006. Although some differences are noted (i.e., more clouds over Madagascar and the Arabian Peninsula), the overall agreement is good, even over land, confirming the statistical results.

e. Information content analysis of the microwave observations

All AMSU-A and -B channels have been used so far in this study to classify the clouds. What are the respective

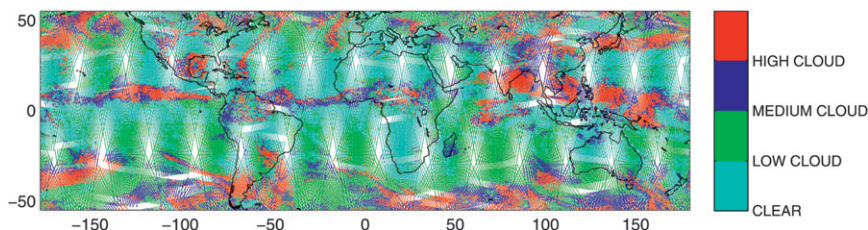


FIG. 9. Extension of the MW-derived cloud classification over the globe on 1 Jul 2006.

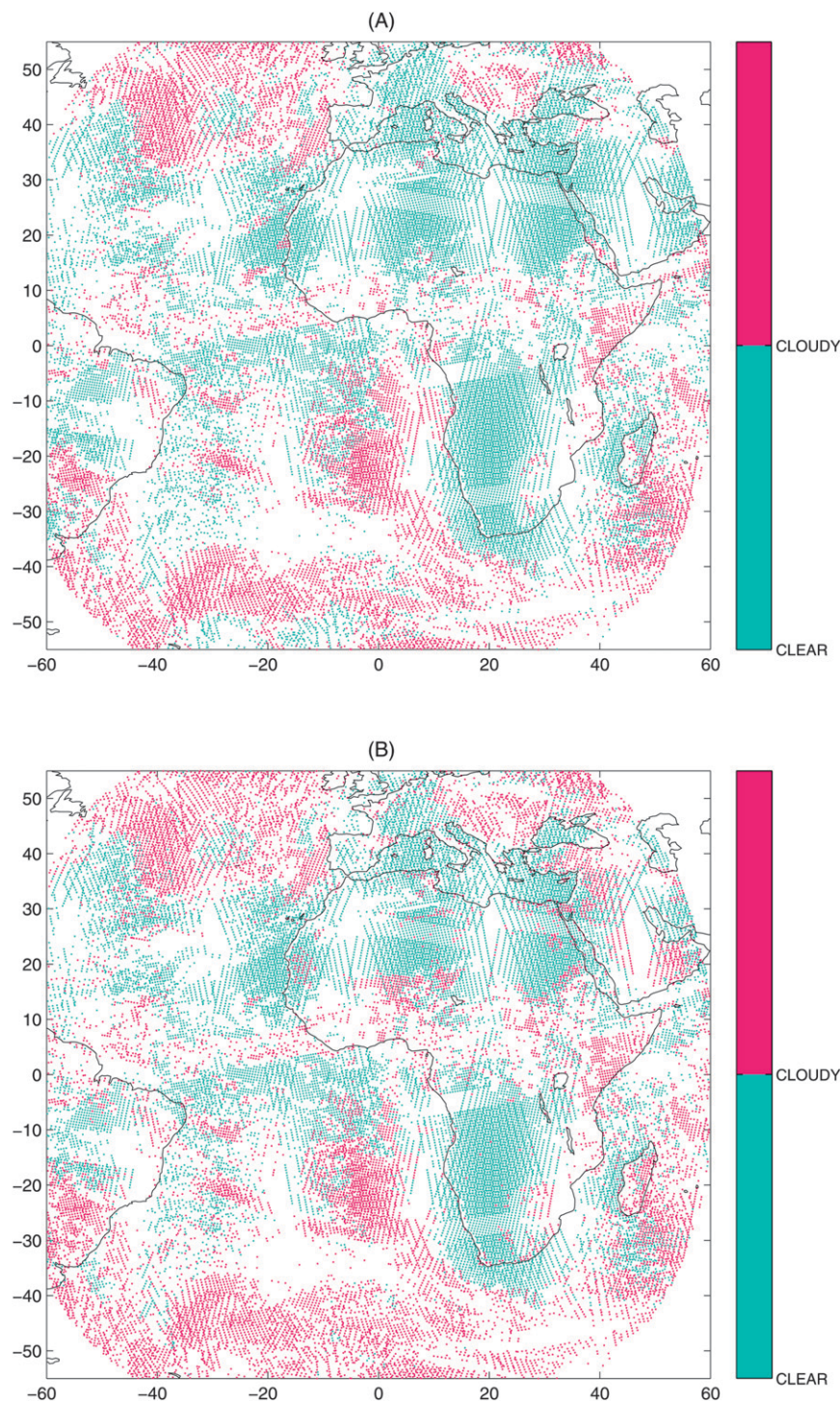


FIG. 10. (a) Cloud mask from SEVIRI and (b) cloud mask from AMSU-A and -B on 1 Jul 2006.

weights of the different channels in the classification results? Which microwave information is the most relevant for each cloud type, the window channels, the temperature, or the water vapor sounding channels?

To answer these questions, we trained the classification model for four configurations: using AMSU-A and -B as before, using only AMSU-A, using only AMSU-B, or using only the window channels. Window channels are

TABLE 6. Cloud classification confusion matrices for the neural network over land and ocean when using AMSU-A plus AMSU-B observations, AMSU-A only, AMSU-B only, and window channels only. These statistics are performed on relatively homogeneous pixels (see section 2c). Boldface signify the diagonal line of the confusion matrices.

Inputs	%	Land				Ocean			
		Clear	Low	Medium	High	Clear	Low	Medium	High
AMSU-A plus	Clear	81.87	14.53	3.06	0.53	87.99	7.46	2.61	1.94
AMSU-B	Low	13.31	73.67	12.23	0.79	9.93	79.91	9.54	0.62
	Medium	3.15	10.09	80.71	6.05	1.70	9.48	77.56	11.26
	High	3.02	7.62	7.62	86.94	2.12	2.69	11.13	84.07
AMSU-A	Clear	68.95	15.38	11.14	4.53	80.95	10.71	2.88	5.45
	Low	14.19	69.40	15.24	1.18	12.44	72.35	10.92	4.29
	Medium	9.97	16.83	63.64	9.56	4.94	11.67	60.25	23.14
AMSU-B	High	8.29	7.19	15.77	68.75	7.74	6.48	22.97	62.81
	Clear	74.93	21.55	2.65	0.87	77.94	16.40	2.85	2.81
	Low	21.73	58.43	18.42	1.42	20.14	66.99	11.43	1.44
Window channels	Medium	2.48	15.51	74.58	7.43	3.34	11.11	70.92	14.63
	High	3.56	3.90	9.85	82.69	3.15	2.62	12.80	81.44
	Clear	73.44	20.35	5.42	0.79	89.41	6.57	1.73	2.28
Window channels	Low	19.45	65.66	14.59	0.30	11.17	76.91	10.12	1.80
	Medium	5.77	13.30	73.83	7.09	3.35	14.23	65.78	16.64
	High	4.04	4.48	22.32	69.16	4.90	4.51	27.57	63.02
No. of pixels		1 474 280	80 606	45 998	85 022	724 979	528 995	47 128	114 020

here AMSU-A 23.8, 31.4, 50.3, 89, and 150 GHz, and AMSU-B channels 89 and 183 GHz. The interested reader can also test all minus one channel to truly measure the information content of each channel. A very interesting test would also consist of using four mutually exclusive subgroups—1) AMSU-A low-frequency window channels (excluding 89 GHz), 2) AMSU-A temperature sounding channels, 3) AMSU-B high-frequency window channels, and (4) AMSU-B water vapor sounding channels—because the window and sounding channels from the two instruments respond to different atmospheric and/or surface signals. This will be the subject of a future study.

Table 6 compares the confusion matrices for these four configurations, for ocean and land separately. There is still one specialized classifier for each surface type but the statistics are given for all the land pixels to simplify the presentation of the results. AMSU-A and -B together provide better statistics than most of these observations. An exception is the better retrieval of clear cases over the ocean using only the window channels. However, it should be noted that when comparing two classifier, it is misleading to look at the detection level of only one class, without analyzing what happens to the other classes.² Overall, the methodology capitalizes on the synergy between the two instruments. Over ocean, window channels play a key role in the discrimination of

the clear cases and in the determination of low cloud cover. For medium and high clouds, the information comes mostly from AMSU-B, regardless of the surface. This could be expected from our sensitivity analysis (section 2e). Over land, clear cases are better isolated by high-frequency observations (AMSU-B). AMSU-A, especially over land, does not provide as much information as AMSU-B, except for a better retrieval of low cloud cases³; over ocean it does exceed the potential of AMSU-B for discriminating between clear cases and low clouds, but only with a limited margin.

5. Conclusions

A statistical cloud classification and a cloud mask have been developed based on AMSU-A and -B observations. The visible and infrared data from MSG-SEVIRI have been used to train the microwave classifier. Clear sky and low, medium, and high clouds can be retrieved over ocean and land at more than an 80% confidence level, using a neural network classification method when applied to relatively homogeneous cloud pixels. These results are very encouraging. Thin clouds have been suppressed from the study and microwave observations are not sensitive to this type of clouds (there is no need to retrieve this type of cloud information here since the forthcoming MW retrieval of surface or atmospheric

² An extreme case would be a classifier that will always predict one class; its level of retrieval of this particular class would be 100%, but the level of prediction of the other classes would be 0%.

³ This statement needs to be balanced as low cloud cases are favored to the detriment of the retrieval of other cloud cases.

parameters is not sensitive to this type of clouds), but the microwave characterization of thicker clouds seems to be possible. The microwave observations have a coarser horizontal resolution than do those from SEVIRI, and this explains part of the discrepancy. When the cloud classification from SEVIRI is heterogeneous inside the bigger MW pixel, the results are degraded, but these errors are related not only to the misclassification but also to the spatial variability inside the MW pixels. It has been shown that using the a posteriori probabilities estimated by the classifier reduces the ambiguities in the classification, but limits the number of scenes to be processed.

As could be expected, the major difficulty is the classification of thin cirrus and partially cloud-covered pixels. They are labeled by the MW classifier as clear but also low and midlevel cloud, especially over land. The low clouds are difficult to isolate, especially over land, because of the contamination of the information by the surface contribution. AMSU-B provides significant information over both land and ocean, especially for the classification of medium and high clouds, whereas AMSU-A is more efficient over ocean when discriminating between clear situations and low clouds.

In the retrieval of atmospheric and surface parameters from microwave observations, an initial important step is often the cloud detection, to avoid any retrieval in cloudy areas because of their contaminating effects or to adapt the retrieval methodology to the presence of clouds. In operational numerical weather prediction centers, external cloud flags from visible and/or infrared sources are not used for the assimilation of microwave observations. In this study, a stand-alone microwave cloud classifier has been developed. It provides reliable results and could be easily adapted to the particular application under study. The microwave classifier cannot reach the level of precision or detail of the SEVIRI-based algorithm, in particular for complex cloud scenes such as partially cloud-covered pixels or multilayer clouds. The important point is that if the microwave observations are not sensitive enough to the detailed cloud characteristics, then it is not worth taking it into account in the cloud classifier because this cloud information has less impact on the satellite measurements, with a reduced corruption of the retrievals.

This microwave cloud classification can be applied globally. The SEVIRI cloud classification algorithm of the NWPSAF (section 2a) is being adapted to other geostationary satellites to obtain global coverage. Since the other visible and infrared instruments on board geostationary satellites have fewer channels than SEVIRI, they cannot offer the same precision in the cloud classification. The microwave global cloud classification derived from SEVIRI could help evaluate the other geostationary algorithms (even if other VIS-IR cloud

classification channels on board polar-orbiting satellites would be a better choice). The confusion matrices should be similar for all geostationary satellites, thus using the microwave classification as a diagnostic tool.

The cloud classification defined in this study can be used to perform MW retrievals of temperature and water vapor over the clouds. Another possibility would be to use the infrared and MW synergy. Since both spectral domains have a cloud information content (e.g., the microwave observations can be used to estimate the cloud liquid water path), using both of them simultaneously would greatly improve cloud characterization. Such a synergetic approach was used for example in Minnis et al. (2007) for ice cloud properties in ice-over-water cloud systems.

Acknowledgments. The authors thank the Institut National des Sciences de l'Univers (INSU) from the Centre National de la Recherche Scientifique (CNRS) and the French space agency, Centre National des Etudes Spatiales (CNES), to support part of this study. This work is a preliminary study for the Megha-Tropiques mission and the authors thank the Megha-Tropiques team and in particular Rémy Roca. Geneviève Sèze would like to thank Météo-France Lannion for providing the SEVIRI cloud classification algorithm. Filipe Aires would also like to thank Alope Mathur from the Indian Space Research Organisation (ISRO), who spent 6 months at LMD and participated in the collocation of the data.

REFERENCES

- Ackerman, S., K. Strabala, W. Menzel, R. Frey, C. Moeller, and L. Gumley, 1998: Discriminating clear sky from clouds with MODIS. *J. Geophys. Res.*, **103** (D24), 32 141–32 157.
- Aires, F., C. Prigent, W. Rossow, and M. Rothstein, 2001: A new neural network approach including first-guess for retrieval of atmospheric water vapour, cloud liquid water path, surface temperature and emissivities over land from satellite microwave observations. *J. Geophys. Res.*, **106** (D14), 14 887–14 907.
- Alishouse, J., S. Snyder, V. Jennifer, and R. Ferraro, 1990: Determination of oceanic total precipitable water from the SSM/I. *IEEE Trans. Geosci. Remote Sens.*, **28**, 811–816.
- Ambroise, C., G. Sze, F. Badran, and S. Thiria, 2000: Hierarchical clustering of self-organizing maps for cloud classification. *Neurocomputing*, **30**, 47–52.
- Baum, B., V. Tovinkere, J. Titlow, and R. Welch, 1997: Automated cloud classification of global AVHRR data using a fuzzy logic approach. *J. Appl. Meteor.*, **36**, 1519–1540.
- Bishop, C., 1996: *Neural Networks for Pattern Recognition*. Oxford University Press, 482 pp.
- Coakley, J., 1983: Properties of multilayered cloud systems from satellite imagery. *J. Geophys. Res.*, **88** (C15), 10 818–10 828.
- Cybenko, G., 1989: Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.*, **2**, 303–314.

- Derrien, M., and H. L. Gléau, 2005: MSG/SEVIRI cloud mask and type from SAFNWC. *Int. J. Remote Sens.*, **26**, 4707–4732.
- , and —, 2010: Improvement of cloud detection near sunrise and sunset by temporal-differencing and region-growing techniques with real-time SEVIRI. *Int. J. Remote Sens.*, **31**, 1765–1780.
- Desbois, M., G. Sèze, and G. Szejwach, 1982: Automatic classification of clouds on Meteosat imagery: Application to high-level clouds. *J. Appl. Meteor.*, **21**, 401–412.
- Engelen, R., and G. Stephens, 1999: Characterization of water vapor retrievals from infrared TOVS radiances and microwave SSM/T-2 radiances. *Quart. J. Roy. Meteor. Soc.*, **125**, 331–351.
- Ferraro, R., F. Weng, N. Grody, and A. Basist, 1996: An eight-year (1987–1994) time series of rainfall, clouds, water vapor, snow cover, and sea ice derived from SSM/I measurements. *Bull. Amer. Meteor. Soc.*, **77**, 891–905.
- Frey, R., S. Ackerman, Y. Liu, K. Strabala, H. Zhang, J. Key, and X. Wang, 2008: Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. *J. Atmos. Oceanic Technol.*, **25**, 1057–1072.
- Govaerts, Y., A. Arriaga, and J. Schmetz, 2001: Operational vicarious calibration of the MSG/SEVIRI solar channels. *Adv. Space Res.*, **28**, 21–30, doi:10.1016/S0273-1177(01)00269-1.
- Greenwald, T., and S. Christopher, 2002: Effect of cold clouds on satellite measurements near 183 GHz. *J. Geophys. Res.*, **107**, 4170, doi:10.1029/2000JD000258.
- , G. Stephens, T. V. Haar, and D. Jackson, 1993: A physical retrieval of cloud liquid water over the global oceans using Special Sensor Microwave/Imager (SSM/I) observations. *J. Geophys. Res.*, **98** (D10), 18 471–18 488.
- Holz, R., S. Ackerman, F. Nagle, R. Frey, S. Dutcher, R. Kuehn, M. Vaughan, and B. Baum, 2008: Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP. *J. Geophys. Res.*, **113**, D00A19, doi:10.1029/2008JD009837.
- Hong, F., and G. Heygster, 2005: Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements. *J. Geophys. Res.*, **110**, D05205, doi:10.1029/2004JD004949.
- Hornik, K., M. Stinchcombe, and H. White, 1989: Multilayer feedforward networks are universal approximators. *Neural Networks*, **2**, 359–366.
- Inoue, T., 1985: On the temperature and effective emissivity determination of semitransparent cirrus clouds by bi-spectral measurements in the 10 mm window region. *J. Meteor. Soc. Japan*, **63**, 88–99.
- Jin, Y., and W. Rossow, 1997: Detection of cirrus overlapping low-level clouds. *J. Geophys. Res.*, **102**, 1727–1737.
- Krzanowski, W. J., 1988: *Principles of Multivariate Analysis*. Clarendon Press, 563 pp.
- Miller, S., and W. Emery, 1997: An automated neural network cloud classifier for use over land and ocean surfaces. *J. Appl. Meteor.*, **36**, 1346–1362.
- Minnis, P., J. Huang, B. Lin, Y. Yi, R. Arduini, T.-F. Fan, J. Ayers, and G. Mace, 2007: Ice cloud properties in ice-over-water cloud systems using Tropical Rainfall Measuring Mission (TRMM) visible and infrared scanner and TRMM Microwave Imager data. *J. Geophys. Res.*, **112**, D06206, doi:10.1029/2006JD007626.
- O'Dell, C., J. Wentz, and R. Bennartz, 2008: Cloud liquid water path from satellite-based passive microwave observations: A new climatology over the global oceans. *J. Climate*, **21**, 1721–1739.
- Prigent, C., F. Aires, and W. Rossow, 2001: Joint characterization of the vegetation by satellite observations from visible to microwave wavelengths: A sensitivity analysis. *J. Geophys. Res.*, **106**, 20 665–20 685.
- , —, and —, 2006: Land surface microwave emissivities over the globe for a decade. *Bull. Amer. Meteor. Soc.*, **87**, 1572–1584.
- , F. Papa, F. Aires, W. Rossow, and E. Matthews, 2007: Global inundation dynamics inferred from multiple satellite observations, 1993–2000. *J. Geophys. Res.*, **112**, D12107, doi:10.1029/2006JD007847.
- , E. Jaumouille, F. Chevallier, and F. Aires, 2008: A parameterization of the microwave land surface emissivity between 19 and 100 GHz, anchored to satellite-derived estimates. *IEEE Trans. Geosci. Remote Sens.*, **46**, 344–352.
- Radnoti, G., P. Bauer, A. McNally, C. Cardinali, S. Healy, and P. de Rosnay, 2010: ECMWF study on the impact of future developments of the space-based observing system on numerical weather prediction. ECMWF Tech. Memo. 638, 115 pp.
- Rossow, W., and R. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.*, **80**, 2261–2287.
- , and Coauthors, 1985: ISCCP Cloud Algorithm Intercomparison. *J. Climate Appl. Meteor.*, **24**, 877–903.
- Rumelhart, D., G. Hinton, and R. Williams, 1986: Learning internal representations by error propagation. *Parallel Distributed Processing*, D. Rumelhart and J. McClelland, Eds., Vol. I, The MIT Press, 318–362.
- Saunders, R., 1986: An automated scheme for the removal of cloud contamination from AVHRR radiances over western Europe. *Int. J. Remote Sens.*, **7**, 867–886.
- Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to the Meteosat Second Generation (MSG). *Bull. Amer. Meteor. Soc.*, **83**, 977–992.
- Seber, G., 1984: *Multivariate Observations*. J. Wiley and Sons, 686 pp.
- Sèze, G., and H. Pawlowska, 2001: Cloud cover analysis with Meteosat-5 during INDOEX. *J. Geophys. Res.*, **106** (D22), 28 415–28 426.
- , H. L. Gléau, M. Derrien, B. Six, and J. Pelon, 2009: Evaluation of the global cloud cover parameters obtained from geostationary data in the frame of the Megha-Tropiques mission with CALIPSO lidar observations. *Proc. 2009 EUMETSAT Meteorological Satellite Conf.*, Bath, United Kingdom, EUMETSAT. [Available online at http://www.eumetsat.int/Home/Main/AboutEUMETSAT/Publications/ConferenceandWorkshop/Proceedings/SP_201001515267179?l=en.]
- Weng, F., L. Zhao, R. Ferraro, G. Poe, X. Li, and N. Grody, 2003: Advanced microwave sounding unit cloud and precipitation algorithms. *Radio Sci.*, **38**, 8086–8096.